

28464

Paper Prepared for the
Canadian Evaluation Society's
5th Annual Conference in
Ottawa, Ontario



LINKING PROGRAM EVALUATION TO POLICY AND
MANAGEMENT INFORMATION NEEDS

March 1984

Eleanor Chelimsky
Director, GAO's
Program Evaluation &
Methodology Division

Note: The views and opinions expressed by the author are her own and should not be construed to be the policy or position of the General Accounting Office.

029071

The Evaluation/Use Linkage

If program evaluation, like basic research, existed only to improve knowledge in a particular area, its success could be measured exclusively in terms of the degree to which it had advanced understanding in that area. But although one purpose of program evaluation is indeed to contribute to such knowledge and understanding, another equally important purpose is to be useful in bringing information to bear on public program and policy decisions of several different types. While there is nothing new in this statement, it does mean that measuring evaluation success must not only include judgments about the degree of increase or improvement in knowledge brought by the evaluation, but also about the degree to which the evaluation was valuable to a particular user or audience in meeting a specific information need. Since it is obvious that all types of knowledge-producing evaluations may not be relevant to all audiences, this raises the question of how program evaluation can best be linked to the differing information needs of its various users.

But to raise this question is simultaneously to raise a subset of component questions. For example, what kinds of information needs can an evaluation be expected to supply? And how do these needs differ for various users, or target audiences? Who are the typical users of an evaluation and what kinds of relationships, if any, exist among them? Should an evaluator, in designing a study, seek to bring information to one user only, or can many users be targeted by the same evaluation? How do new programs differ from existing programs, and controversial programs from accepted programs, in their evaluative information needs? Are all types of

evaluations likely to be useful at different times to different audiences, or are some information needs unique to particular functions, particular audiences, particular times, so that certain types of evaluations will be predictably more relevant to some users than to others? In sum, given a variety of demands -- that is, information needs stemming from a variety of users performing a variety of public functions -- and given also a variety in supply -- that is, the different program evaluation types available to meet those different needs -- how are demand and supply currently linked? How should they be linked? And, most importantly, how can that linkage be improved?

This paper proposes to examine the linkage between program evaluation and its users at the federal level. In order to do so, it will explore the relationship between different types of functions and users, different types of information needs, different times in the life of a public program, and different types of evaluation studies. It will focus, not upon the qualities and failures of the evaluation types or upon the merits of the information needs, but upon the evaluation/use linkage itself. It will look both at the effective demand for evaluative information and at the supply of program evaluation types available to meet that demand. The purpose is not only to improve the targeting of program evaluations upon general information needs, but also to improve the targeting of specific types of evaluations upon specific types of information needs, and, as a result, to improve the usefulness of evaluation findings for policy or program management decisionmaking, which is, of course, the subject of this Canadian evaluation conference.

Because the author's experience is largely with public programs in the United States, the framework employed here is based on relationships existing in the American Government. But although parliamentary and presidential types of government are not the same, and although basic policymaking and program management functions are distributed differently in the two governmental cases, the functions themselves and the information needs deriving from them are sufficiently similar to make much of the analysis applicable in both cases.

What Are the Purposes That Evaluation May Serve?

Program evaluations serve both general audiences (such as the public or the media which are normally the ultimate users of many evaluations) and individual public decisionmakers with particular information needs. These decisionmakers may be in the executive or legislative branches of government, they may play management or policy roles with respect to public programs, and they may need information from evaluation for three very broad kinds of purposes:

- for policy formulation -- that is, with regard to public programs, to assess and/or justify the need for a new program;
- for policy execution -- that is, to ensure that a program is implemented in the most cost/effective way; and
- for accountability in public decisionmaking -- that is, to determine the effectiveness of an operating program and the need for its continuation, modification, or termination.

The purpose of policy formulation, as it applies to new programs, requires information from evaluation in at least three major areas:

(1) information on the problem or the threat addressed by the program; (2) information on the results of past programs or related efforts that attempted to deal with the problem or threat; and (3) information allowing the selection of one alternative program over another. The evaluative information required for policy execution, as it applies either to new or to existing programs, is quite different from that required for policy formulation. It includes: (1) information on program implementation (such as the degree to which the program is operational, how similar it is across sites, whether it conforms to the policies and expectations formulated, how much it costs, how stakeholders feel about it, whether there are major problems of service delivery or error, fraud and abuse, etc.); (2) information on program management (such as the degree of control over expenditures, the qualifications and credentials of personnel, the allocation of resources, the use of program information in decisionmaking, etc.); and (3) ongoing information on the current state of the problem addressed by the program. The purpose of accountability, as it applies to both new and existing programs, requires information from evaluation which again differs markedly from that required for the other two policy purposes. Here evaluations must emphasize: (1) information on program outcomes; (2) information on the degree to which the program made, or is making, a difference (that is, the program's impact on the problem or threat addressed); and (3) information on the unexpected (as well as the expected) effects of the program. These different types of evaluative information needs that are related to the three kinds of general policy purposes have implications, of course, for the types

of evaluations that are most appropriate to the particular information need. This point will be discussed later on in some detail.

What Are the Typical Functions or Roles of Evaluation Users and How Do They Differ?

Who, then, are the typical governmental users of evaluative information developed for the three policy purposes, how do their functions or roles differ, and to whom are they accountable? Among the types of executive and legislative branch decisionmakers in the United States who need and use evaluative information to serve one or more of these purposes are the following six:

In the executive branch

- program managers,
- agency heads and top policymakers in a given agency or department, and
- central government budget and policy authorities.

In the legislative branch

- the four legislative agencies, that is, the Congressional Research Service (CRS), the General Accounting Office (GAO), the Office of Technology Assessment (OTA), and the Congressional Budget Office (CBO),
- authorization, appropriations, and budget committees, and
- oversight committees.

Executive branch program managers seek evaluative information to plan, implement, manage, and eventually modify their programs so as to make them as cost/effective as possible. Their accountability for this work is essentially to the head of their agency or

to its top policymakers. The functions and roles of the latter, however, require somewhat different types of evaluative information than do those of program managers. For example, a departmental secretary is less interested than a program manager in detailed program information but will instead need data to determine and/or justify the need for a new program, to assess the effectiveness of an operating program, or to review the need to continue or modify a program. The accountability of an agency's top managers is both to the Administration (i.e., to the Office of Management and Budget (OMB) as well as to the policymaking Executive Office of the President) and to the Congress (through its authorization, appropriations, budget, and oversight committees). With regard to staff from OMB and the Executive Office, these decisionmakers -- in their central budget and policy functions -- may also look to evaluation to inform them on the effectiveness of a program or to help in the decision to continue, cut or otherwise change it. Their accountability is to the President.

In the legislative branch, Congressional authorization, appropriations, and budget committees use evaluation findings essentially as contributing information to program funding or refunding decisions. Congressional oversight committees, on the other hand, rely on evaluation not only to supply findings about agency programs but also to bring information about how agencies are performing their various functions (such as personnel and resource management for example, or the manner in which an agency has organized itself to conduct evaluation or to ensure internal

controls). Although all four of these types of congressional committees may use executive agency-produced evaluation findings in their negotiation and decisionmaking processes, they also rely on evaluation findings from independent sources such as universities, consulting firms and other groups. They may use those findings directly, through the work of congressional staff, or indirectly, through the proxy of one or another of the four legislative agencies. CRS may, for example, report findings to congressional sponsors from a wide variety of evaluations in an "issues" paper for legislative use. The GAO, which has the congressional mandate for performing program evaluations, may either report its own findings to the Congress, or analyze evaluation findings from many different sources to inform the Congress, or critique methodological or technical aspects of existing evaluations when findings conflict (or appear to conflict). OTA may use evaluation findings as a foundation from which to forecast technological impacts, and CBO may use them as part of the empirical base on which to construct the likely outcomes of alternative economic policies. The accountability of the legislative agencies is to the Congress; the Congress' accountability, like that of the President, is, of course, to the public via regular general elections.

The six different types of evaluation users, then, have some purposes in common and some that are uniquely their own. As a result, there are some areas of evaluative information need that they share and some that are distinct. Only a program manager or the GAO, for example, may need to receive fine-grained information

on implementation and operational issues specific to one or two local sites in a large national program. Yet, even though there may not be many audiences for this information, it may be crucial to the needs of those two users. On the other hand, evaluation findings on:

- the effectiveness of a program,
- client satisfaction or frustration with program services,
- the views of program stakeholders,
- trends in the problem(s) addressed by the program, or
- changes in the dimensions or focus of the program

are likely to be helpful to all users, and may interest general audiences as well.

However, the degree of detail needed by users from evaluative information which all of them may want, will differ depending upon the user's specific function, purpose, and particular accountability. A Member of Congress, for example, may need to be aware only of the major findings of a program evaluation, whereas an agency head -- who may be called upon by OMB, the President, or the Congress to defend the program or explain the findings -- will require a much more detailed knowledge of the information produced. This means that although the basic information produced by an evaluation may be useful to several different audiences, different versions of that information, presenting different levels of detail, may be needed.

The question of the appropriateness of a particular evaluation to multiple user needs is one that evaluators ought to consider very carefully. Evaluations are so few in number, and the idea of

serving several users with the same evaluation is so attractive that a danger exists of introducing information distortion inadvertently. Explicit statements by the evaluators about the objectives and limitations of the evaluation are needed to ensure that an evaluation performed, say, for an agency head's information, to learn the opinions of practitioners or stakeholders about a program is not interpreted by a secondary user -- the press, for example -- as an evaluation of program effectiveness.

Further, the ability to serve multiple users is always constrained by the need for an evaluation to answer the precise question(s) posed by the primary user, and the need for the conclusiveness of the answer(s) given to match that user's particular information need. For example, it is reasonable to undertake an expensive, large-scale effectiveness evaluation when the primary user must have the most conclusive information possible. But if an evaluation user's objective is to negotiate rather than to legislate, or to develop some general information in an area where little currently exists, then less conclusive -- and less costly -- information may be both appropriate and desirable.

Finally, it is probably important to point out an essential difference between evaluative information designed for executive branch and legislative branch users. While both branches need evaluative evidence on, say, the need for a program as part of their policy-formulating role (that is, their planning or authorizing functions), and while they both need evidence of efficiency and effectiveness to support their accountability role in their administrative and legislative oversight processes, it is the

Administration which is charged with policy execution and the Administration, therefore, which actually implements programs. As a result, executive branch program managers and -- to a lesser degree -- agency heads, may often (and quite properly) be virtually the sole users of certain types of detailed evaluative information which supports decisionmaking about program operations.

This suggests that, as a general rule, evaluators, in thinking about the eventual use of their evaluations, can expect their work on program effectiveness and on the feasibility and logic of proposed new programs to be more valuable to a greater variety of secondary users than their work on program operations.

How Do Evaluative Information Needs Differ
for New Programs Versus Old Ones? For
Accepted Programs Versus Controversial Programs?

Information needs are not only different for different users; they also differ as between new and existing programs. A new program has a need for planning information that an existing program typically does not present. But an existing program needs effectiveness and other information on a periodic basis that can only be obtained when a program has been fully operational for some time.

Information needs are also affected by a program's dimensions, life cycle, and by the circumstances surrounding it. In a new program, for example, it makes a difference, with regard to the evidence required to justify its need, if the program is large or small, expensive or modest, controversial or consensual, presenting many unknowns or instead, clearly feasible. A large and expensive program needs stronger evaluative justification than does a small,

modest one, because of accountability to the taxpayer. This is often important for defense programs (like the MX missile and B1 bomber, for example) which cost a great deal and whose logic and assumptions are typically subject to extensive debate, first within the Administration, and then, by the Congress, and in the press.

Even when new programs are small, however, they need to have a very sound basis and rationale if they are controversial. A recent example was the Reagan Administration's proposal to end the moratorium on the production of nerve gas and to begin augmenting the existing stockpile of unitary weapons with new "binary" weapons. The appropriations requested were not large, as defense programs go. However, emotions surrounding the issue of chemical warfare run high, the moratorium had existed for many years, the Administration's case for ending the moratorium and starting production on binary weapons was not supported by strong evidence, and the Congress rejected production of any new chemical weapons.

Proposed programs whose scientific or technological feasibility is still unclear also need to have sound logic and convincing presumptions of effectiveness behind them. When feasibility problems are real, they are sure to surface in the debate on funding, and should therefore be persuasively addressed in the program rationale. Yet it sometimes happens that an agency may believe a new research program is needed, and that its feasibility problems cannot be resolved without federal funds. Under those circumstances, the temptation can be very great to minimize the feasibility problems so as to obtain appropriations for the program. While this strategy may work over the short term, it tends to mislead policymakers, to create overoptimistic expectations for

the program, and to give credence to a highly simplistic and mythical view of the research process. In particular, by minimizing technical problems, the implication is that:

(i) research regularly uncovers new ideas that (ii) are then fed into the hands of development engineers who (iii) neatly establish the technical and economic feasibility of a new gadget or technology which (iv) can then be introduced smoothly into efficient production, and thus (v) the research-to-production process effortlessly solves some problem or meets some national need. (Nichols, Science, 1971)

Of course, the long-term effect of this oversimplification of the research process -- which by its nature is iterative, non-linear, creative, and highly serendipitous -- will be the quasi-certainty of negative evaluation findings if the program's evaluation should mistakenly structure its criteria for success on the agency's rhetoric. This is because the presence of significant technical unknowns in a program means that events are likely to be unpredictable and that major problems may arise in development or testing that could cost a lot of money and take a long time to solve.

The point here is that oversimplifying and overpromising about what evidence is available to rationalize funding for a new program can misfire badly. One result just discussed is that evaluation findings for the program may be meaningless if the program's objectives were overstated and unachievable from the start. Another result is that efforts to deal with the evaluation failures (e.g., by modifying objectives, continuing to test and retest, etc.) can cause a further stretching out of what may already have been a long development cycle. A third -- and perhaps the most important -- result is that the credibility of the overpromisers inevitably suffers over the long term as, one after the other, the

negative evaluation findings come in on programs about which so very much had been promised.

All of this underscores the truly critical need for evaluative information -- and for an evaluative mechanism capable of intervening -- in the policy-formulation process that generates new programs. This is especially important for new programs that are controversial and for those that present significant scientific or technical unknowns.

How Do the Purposes, Functions and
Accountabilities of Evaluation Users
Shape Their Specific Information Needs
Vis-a-Vis New and Existing Programs?

Given six types of evaluation users with differing functions, accountabilities, and general information needs vis-a-vis public programs, what can be said about their specific information needs? Are these quite distinct from each other or are there areas of commonality? Table I lists examples of specific evaluative information needs each user might be expected to have, based on purpose, function, accountability, and on whether or not the program is an existing one or a new one.

Each of the examples of information need shown in the table is directly related to the functions or roles of a particular user. Thus, the likely relevance to other users of each piece of evaluative information generated in support of an information need can be tracked across the columns of the table. For example, it is clear from Table I that evaluative information which can be used to monitor and analyze the problem or threat addressed by an existing program is not only useful to the program manager, but also to the

SIX EVALUATION USERS

TYPE OF PROGRAM	(1) <u>USER</u> Executive Branch Program Manager <u>FUNCTION</u> Policy Execution Program Planning Implementation and Effectiveness <u>ACCOUNTABILITY</u> To Agency Head or Top Agency Policymakers	(2) <u>USER</u> Executive Branch Agency Head Top Agency Policymaker <u>FUNCTION</u> Policy Formulation and Execution Administrative Program Oversight <u>ACCOUNTABILITY</u> To Central Budgetary and Policy Authorities to the President and to the Congress	(3) <u>USER</u> Executive Branch Central Budgetary and Policy Authorities <u>FUNCTION</u> Policy Formulation and Oversight of Policy and Program Execution <u>ACCOUNTABILITY</u> To the President	(4) <u>USER</u> Legislative Branch Agencies (FRS GAO OIA and CBO) <u>FUNCTION</u> Policy Formulation and Oversight Budget Analysis Audit and Evaluation Support to the Congress <u>ACCOUNTABILITY</u> To the Congress	(5) <u>USER</u> Legislative Branch Authorization, Appropriations and Budget Committees <u>FUNCTION</u> Policy Formulation and Oversight Enactment of Priority Programs Within Budget Constraints <u>ACCOUNTABILITY</u> To the Public	(6) <u>USER</u> Legislative Branch Oversight Committees <u>FUNCTION</u> Legislative Oversight of Executive Branch Programs and Management <u>ACCOUNTABILITY</u> To the Public
NEW PROGRAM PROPOSAL	<u>INFORMATION NEEDS</u> <u>Program Planning</u> For Example Monitoring and analysis of the problem threat Effectiveness of past efforts to solve the problem, or related problem or threat Comparative costs and benefits of alternative program options Logic of relationship between problem threat and program assumptions Determination of likely feasibility, costs, acceptability and effectiveness vis a vis problem <u>Assurance of Future Information</u> For Example Pilot test Evaluation design Internal controls Audit plan Data system design to monitor problem program	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Results of program manager's analysis Critique of program planning effort Comparative costs, benefits, and budgetary impact vis a vis other agency programs and general administration policy Constituency stakeholder, and legislative acceptability of program Continued monitoring of problem threat <u>Evaluation Decisions</u> For Example Kinds of future information to be developed In house grant or contract evaluation approval	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Results of program manager's analysis Strength of administration constituency legislative support for program Comparative costs and benefits across agencies and programs vis a vis administration policy goals and budget constraints Examination of degree of duplication, overlap, goal conflicts, etc between proposed new program and existing programs <u>Evaluation Decisions</u> For Example Review of agency head decisions Modification to kinds of information being developed	<u>INFORMATION NEEDS</u> <u>Review of Program Planning</u> For Example Examination of evidence presented in support of program Verification of alternatives to program considered Assessment of program logic related to problem or threat Reanalysis of likely feasibility costs, benefits, and budgetary impact of program Examination of controls against error, fraud and abuse Review of evaluation and data system design <u>Recommendations</u> For Example Other options Evaluation audit language	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Results of executive branch analyses (program proposal and rationale) Results of analyses by legislative branch agencies Review by committee staff Hearings/debate Assessment by committees of strength of constituency/congressional support for program versus committee priorities and budget constraints <u>Evaluation Decisions</u> For Example Audit, evaluation language Modifications to program or evaluation plans Data collection	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Assessment of oversight needs for the program Review of legislative agencies' analyses of likely program feasibility, cost and effectiveness Review of plans for controls, audit, evaluation, and management <u>Evaluation Decisions</u> For Example Audit/evaluation language Modifications to program, evaluation, or management plans Data collection
NEWLY IMPLEMENTED OR EXISTING PROGRAM	<u>INFORMATION NEEDS</u> <u>Program Implementation</u> For Example Costs Extent operational Size of population being served compared to expectations Practitioner satisfaction Client satisfaction Degree of error fraud abuse Relationship of program as implemented to legislative and administration intent Results of problem threat monitoring <u>Program Effectiveness</u> For Example Outcomes of the program Impact made by the program on the problem threat Expected unexpected effects <u>Program Decisions</u> For Example Modification Reevaluation	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Monitoring of costs stakeholder satisfaction levels of error fraud or abuse Assessment of program effectiveness findings Determination of continuing need for program or program components Analysis of program's budgetary impact on agency Monitoring of problem threat Program continuation or modification decisions <u>Evaluation Decisions</u> For Example Assessment of information gaps Agenda for development of missing or inadequate information	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Assessment of evaluation audit findings Assessment of agency head's budgetary and policy decisions vis a vis the program Determination of strength of program success as related to administrations, new goals, priorities, problems Reexamination of political constituency needs <u>Evaluation Decisions</u> For Example Review of agency head decisions on information gaps Modifications to agency evaluation agenda	<u>INFORMATION NEEDS</u> <u>Program Review</u> For Example Methodological review of implementation findings Methodological review of effectiveness findings Re analysis of costs and benefits Review of program management Original review of program implementation or effectiveness Analysis of agency decisions Analysis of present and future budgetary impacts Analysis of current trends in the problem threat addressed by the program <u>Recommendations</u> For Example Further analysis of program effects Program modification Evaluation audit or other modification	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Results of executive branch evaluations and audits Results of legislative branch methodological reviews of findings Review by committee staff Assessment by committee of continued need and support for program in view of budget constraints and status of problem/threat Determination of program modifications <u>Evaluation Decisions</u> For Example Mandate to evaluate or audit aspects of the program Changes in program regulations such as data collection requirements etc	<u>INFORMATION NEEDS</u> <u>Program Decisions</u> For Example Analysis of the results of all evaluations, audits, and analyses performed by the executive branch Results of legislative agencies' reanalysis or evaluations Assessment of match between implemented program and legislative intent Examination of agency management of program and of control systems Determination of program/agency modifications <u>Evaluation Decisions</u> For Example Mandate to evaluate or audit program Mandate to review agency evaluation process Mandate to review management process

Table 1 Evaluation Users' Information Needs

agency head, to legislative agencies and to congressional committees of both types. In the same way, for a new program, the program manager's analytical justification of that program will be examined over and over by a bevy of other evaluation users. On the other hand, the information need for comparative evaluation of competing programs in terms of Administration policy goals is likely to be restricted in some measure to OMB and to the White House.

Looking across the columns of Table I, an important distinction to be made among the six user groups is that although all groups may need evaluative information in support of their functions, not all are responsible for producing them. In fact, if evaluation producers are distinguished from evaluation reviewers and users, then the six groups of evaluation users can be collapsed into four, because two groups -- the program managers and legislative agency staff -- are responsible for developing a large part of the evaluative information required. The evaluative work done by these producers (who are, of course, often users as well) is then available for re-analysis and re-use by policymakers in both executive and legislative branches. *

Again, looking at Table I and bearing in mind the three policy purposes discussed earlier -- policy formulation, policy execution,

*The responsibilities for evaluation production and use given here describe current American norms, which have developed in a fairly ad hoc manner. This situation differs in other countries, notably in Canada, where the process has been carefully developed and where the deputy head of an agency has the overall responsibility for producing and using evaluations of agency programs, see the 1983 Report of the Auditor General of Canada on Program Evaluation.

and accountability -- five basic program activities, performed in varying degrees across the spectrum of evaluation users, can be derived. These are:

- (1) planning and rationalizing a program and its evaluation,
- (2) implementing and managing a program,
- (3) justifying the effectiveness of program implementation and management,
- (4) demonstrating the effectiveness of a program, and
- (5) measuring ongoing problem or program progress.

All of these activities call for specific types of evaluative information which, of course, needs to be produced before it can be used or reviewed. But among the users of evaluation, which ones are also producers, and how does that affect the linkages between evaluation and use?

What Is the Distribution of Producers,
Reviewers and Users of Evaluation
in Terms of the Five Program Activities?

Table II examines the five activities listed above in terms of the role of the evaluation user in producing, reviewing, or using the evaluative information generated in support of the activities. For the five activities requiring evaluative information, the table shows considerable congruence with regard to the use -- if not the production -- of the evaluative information by the four groups. For example, knowledge needed by all reviewers and users of a new program proposal (included in Table II, first column on the left) necessarily involves several types of evaluative information, some

<u>Type of User</u>	<u>Five Types of Activities Generating Evaluative Information Needs</u>				
	Planning and Rationalizing a Program and its Evaluation	Implementing and Managing a Program	Justifying the Effectiveness of Program Implementation and Management	Demonstrating the Effectiveness of a Program	Measuring Ongoing Problem/Program Progress
Program Manager	Producer/User	Producer/User	Producer/User	Producer/User	Producer/User
Agency Head/ Central Agencies	Reviewer/User	Reviewer/User	Reviewer/User	Reviewer/User	Reviewer/User
Legislative Agencies	Reviewer/User	Reviewer/User	Producer/Reviewer/ User	Producer/Reviewer/ User	Producer/Reviewer/ User
Congressional Committees	Reviewer/User	Reviewer/User	Reviewer/User	Reviewer/User	Reviewer/User

Table II: Producers, Reviewers and Users of Evaluative Information

of whose dimensions were already outlined in Table I. Program managers, executive branch agency heads and central authorities, as well as legislative agencies and committees, all need, and are accountable for, evidence of the problem or threat that a proposed new program is designed to solve or meet, and evidence that the program is likely to solve that problem or meet that threat. Such evidence can come from evaluations that (1) develop basic data on what is known about the problem being addressed by the new program or about programs of the past which addressed that problem; (2) identify knowledge gaps -- that is, determine what is not known about the problem and about programs of the past to address it -- so as to accurately estimate the likelihood of program success; (3) distinguish the relative merits of past positions taken by theorists, analysts and advocates with regard to the program or policy -- especially conflicting positions; and (4) assess the prior findings of efforts designed to determine the effects of implementing alternative policies or programs.

In a similar way, for an implemented or existing program, the same reviewers and users will require some level of information on the operational realities of the program, on its effectiveness, and on its present and future budgetary impacts.

However, although the generic needs may be the same, the specific information need of each user and the level of analysis requirement may be different. For example, both the Administration and the Congress may need evaluative information on program effectiveness. But while the Administration's need could perhaps be satisfied with evaluation studies that are not generalizable to

the program as a whole, generalizable information might be precisely what a congressional authorizing committee would need to have.

Similarly, a unit of analysis requirement at the state level that might yield sufficient information for an agency head's or congressional committee's problem monitoring activity might not be sufficiently sensitive for a program manager's or legislative agency's analytical needs. In the oversight of a criminal justice program, for example, it might be sufficient for a congressional committee to have state-level data on crime rates. To be useful to program managers needing to understand program activity results and implement changes in the program on a site-by-site basis, however, these data would have to be broken down much further (by type of crime and offender, by season, by time of day, by location, for example). In order to understand why proactive patrol by police has been unsuccessful in, say, shopping malls and to locate it in places where it will be successful in reducing crime, program managers must have information about the places and times that constitute targets of opportunity for particular types of crimes and criminals.

These differences in specific information need and in unit of analysis requirement have major implications for the evaluations that need to be done to satisfy each user and for the degree to which a single evaluation can serve multiple users. In each case, therefore, it is important for an evaluator to be able to recognize and identify the dimensions of each evaluation user's information need and, as a result, to make conscious decisions in the design

phase about what information will or will not be obtained and available to that user.

On the other hand, it seems clear that some evaluations, at least, can address the information needs of many users. For example, a study of the effectiveness of past efforts to solve a problem, which a producer (i.e., program manager or legislative agency) would prepare to justify a proposed new program, could be useful to all evaluation reviewers and users without any change in the content of the analysis.

But which are the types of evaluation that support the specific activities and information needs of evaluation users? Put another way, what kinds of evaluations are now routinely performed and what are the indicators for the relevance to the specific program or policy information needs of individual evaluation users?

How are Specific Types of Evaluation Studies
Linked to Specific Program and Policy
Information Needs?

The Evaluation Research Society (ERS Standards Committee, 1982) has identified six types of routinely conducted evaluation studies. These are:

(1) Front-end analysis

This is evaluative work that is typically done before deciding to move ahead with a new program.

(2) Evaluability assessment

This examines the logic of a program's assumptions and activities in terms of its objectives, describes the characteristics of program implementation, determines the feasibility and usefulness

of performing an evaluation of the program's effectiveness, and, if the latter determination is positive, lays the groundwork for such an evaluation. In that sense, an evaluability assessment may be the first phase of a larger evaluation effort.

(3) Process evaluation

This is a form of evaluation that either stands alone or is developed in combination with another type of evaluation. As a stand-alone, its purpose is usually to analyze the processes of program implementation -- management strategies, operations, costs, interactions among clients and practitioners, error rates and so forth -- so as to improve them. In combination with another evaluation type (most often an effectiveness evaluation), its purposes may include (a) helping to determine the design of the effectiveness evaluation or (b) helping to explain its findings. In the first case, the process evaluation will precede the effectiveness evaluation, in the second, the two will be coordinated more or less simultaneously.

(4) Effectiveness or impact evaluation

This type of evaluation seeks to find out how well a program is working. To do this, it is necessary to be able to show that any changes observed are in fact a result of the program, rather than of other factors or forces. This means that the design for this kind of evaluation needs to include a basis for comparison that permits an understanding of what conditions would have been the absence of the program.

(5) Program and problem monitoring

This type of evaluation tracks progress (long-term or short-term) in the areas of changes in the problem addressed by the pro-

gram, program compliance with policy, service delivery, numbers of clients served, etc. Administrative data systems that can be very useful to all types and many phases of evaluation, often develop around the program and problem monitoring effort.

(6) Metaevaluation or evaluation synthesis

This is a form of evaluation that uses existing evaluations to determine what has been learned about a program. Depending on the availability of evaluations and other empirical work, this type of evaluation can address many different evaluation questions, including those about the effectiveness of the program and about the extent of existing knowledge in a given program or problem area.

To this may be added a seventh type of evaluation which is the case study. This evaluation strategy is less well defined than the ones identified earlier. For the purposes of this paper, it can be called "an analytic description of an event, a process, an institution or a program" (Hoaglin et al, 1982). A case study may use one of the six forms given by the ERS standards, but it always has the special characteristic of yielding rich, in-depth information about an individual instance of a program or process. The evaluative case study can be used as a stand-alone (e.g., to invalidate a conventionally accepted hypothesis); in combination with another evaluation (e.g., to examine how findings from individual cases may relate to national findings); or cumulatively (e.g., to build up evidence piece by piece when a program is so complex or large-scale that evaluation of the whole is infeasible or must be delayed).

When these seven types of program evaluation studies are linked with the five types of activities generating evaluative

information needs (shown in Table II), this gives a general idea of the number and kinds of purposes that program evaluations can serve for the six groups of users identified earlier (see Table I).

Table III indicates some of the possible linkages.

Several points are immediately obvious in examining Table III. First, program and problem monitoring data are useful for every one of the five types of activities. This is because such data can help establish the range and frequency of a problem (before it has even been decided to propose a program), can inform on progress in implementing and managing the program, can be used -- in conjunction with other efforts -- to determine the effectiveness of both the program as a whole (administrative time series data can be critical here) and its implementation or management, and is the instrument of choice for maintaining a minimum level of awareness of ongoing progress in either the program itself or the problem it is intended to address.

A second point is that establishing the effectiveness of a program can involve all seven types of evaluation, some of them used in conjunction with each other. This underscores the difficulty of demonstrating program effectiveness, the usefulness of employing several different methods in the same evaluation to better assure that demonstration, and the frequent costliness of these evaluations. (However, it is also the case that, as discussed earlier, the cost is usually directly related to the conclusiveness needed for the findings: where the information needed does not have to be the firmest possible, an evaluation synthesis alone could suffice. On the other hand, process evaluations,

<u>Type of Activity</u> <u>Generating</u> <u>Evaluative</u> <u>Information</u> <u>Need</u>	<u>Seven Types of Program Evaluations</u>						
	Front end Analysis	Evaluability Assessment	Process Evaluation	Effectiveness or Impact Evaluation	Program and Problem Monitoring	Metaevaluation or Evaluation Synthesis	Case Study Evaluation
Planning and Rationalizing a Program and Its Evaluation	X				X	X	
Implementing and Managing a Program		X	X		X		X
Justifying the Effectiveness of Program Implementation and Management			X		X	X	X
Demonstrating the Effectiveness of a Program	X	X	X	X	X	X	X
Measuring Ongoing Problem/Program Progress	X				X		

Table III. Linkage of Program Evaluation Types with Evaluative Information Needs of User Groups

monitoring, front-end evaluation, and evaluability assessments, used alone, cannot normally supply evidence of effectiveness adequate for, say, the legislative oversight function).

Third, five types of evaluation -- front-end, process, monitoring, evaluation synthesis, and case study -- are quite versatile with regard to their applications to user activities; all have three or more applications. Evaluability assessments and effectiveness evaluations are somewhat more restricted with regard to the roles they can play.

In terms of the three policy purposes discussed earlier, although seven types of evaluation serve the needs of accountability (as reflected in the demonstration of implementation, management or program effectiveness), and four types serve policy execution, only three serve the needs of policy formulation (i.e., planning and rationalizing a program) in the executive branch. This is especially striking in that policy execution (i.e., implementing and managing a program) is the activity of only one type of evaluation user (see Table II) the program manager, while that of policy formulation is common to all three executive branch users. This is not to say that front-end analysis, problem monitoring, and metaevaluation are not powerful tools in support of policy formulation. But given the nearly exclusive production role of program managers in the executive branch (see Table II), it is not apparent how these evaluations -- which could help in major ways to assist the decisionmaking of agency heads, central budgetary and policy authorities -- can be or are being developed. Here, one sees a major advantage of the Canadian system which focuses the

entire authority and responsibility for evaluation (i.e., both for production and for use) on the deputy head of an agency.

With regard to legislative users, the lack of evaluations supporting policy formulation in the executive branch puts the onus on those other evaluation producers, the legislative agencies, to fill the void. Unfortunately, the oversight needs for evaluative information have been so overwhelming, and the legislative agencies have been so understaffed compared to the work they need to accomplish, that legislative information requirements to support either congressional or Administration proposals for new programs -- for example, how needed are they? what is known about past efforts to address them? are they feasible? will they bring the expected results? -- may well be one of the most significant areas needing reinforced emphasis by program evaluators. This takes on added importance as complex proposals for new programs begin to account for more and more of the national budget dollars and as scientists and engineers continue to develop more and more complex technology. This is especially true in the defense sector of the budget. Yet it is precisely in defense programs that explanations, justifications, and supporting data for new programs have not only been in short supply, but have also been presented -- at least since the 1960's -- in terms of highly optimistic and even misleading rationales.

Another reason why evaluators should pay more attention to front-end analysis and to the needs of policy formulation in both branches of government is that the benefits of early problem identification considerably outweigh their cost, especially from

the taxpayer's point of view. While establishing program effects is a very important function for evaluation in support of oversight, it is also costly and often unsuccessful because the program is so far along and its advocates are so well entrenched. This is all the more reason for evaluators to try harder to ensure the presence of sound information early on in the executive and legislative branch debates about new programs.

Summary and Conclusions

This paper has sought to establish a framework by which to link types of evaluation with managers' and policymakers' information needs. This involved, first, identifying three broad policy or program purposes which evaluations may serve (that is, policy formulation, policy execution, and accountability); pinpointing six types of evaluation user groups that generate evaluative requirements; and then characterizing those evaluative information needs with respect to new as well as operational programs. This analysis describes the foundation, or the strategic demand, for evaluation in public policy and decisionmaking at the federal level. In addition, since each example of information need is directly related to the function or role of a particular user, the likely relevance to other users of each piece of evaluative information can thus be tracked.

Next, five categories of activities generating evaluative information needs were derived from the analysis, activities that can be linked to current practice in program evaluation. The six user groups were collapsed into four on the basis of their producer or user role in evaluation, and then the activity categories

requiring information support were examined in terms both of their distribution across groups and the role of the groups with regard to production, review, or use of the needed evaluative information.

The results of this second analysis add more precision to the question of the multiple use of a single evaluation and suggest that many users can benefit from information needed for all the five activities. However, the type of need and the appropriate unit of analysis may be different.

Seven frequently performed types of program evaluation were then identified and linked with the five activities driving evaluative information needs, giving more precise indicators with regard to multiple use of the same evaluation. First, program and problem monitoring information appears to be useful for every one of the five types of activities. Second, five types of evaluation -- front-end, process, monitoring, evaluation synthesis, and case study -- seem quite versatile with regard to their applications to user activities, whereas evaluability assessments and effectiveness evaluations are somewhat more narrowly focused with regard to the roles they can play.

Other points resulting from the analysis are:

- the importance for the evaluator to be familiar with as many types of evaluation as possible when designing efforts to establish program effectiveness;
- the currently better targeting of types of evaluation to accountability and policy execution than to legislative and especially, executive policy formulation; and

- the need for evaluators to devote more effort to working in the policy formulation area, particularly, in analyzing the justifications for new programs and policies.

Evaluation to support policy formulation is especially important when programs are controversial, or involve major technical uncertainty.

Several problems need to be overcome if proposed new programs are to receive better evaluative scrutiny early enough to be useful. First, there is currently little effective demand for such scrutiny and few effective producers of the needed information. Second, timeframes for this kind of information are typically short, and evaluators will need to adjust to that constraint. Finally, there seems to be no institutional executive branch mechanism for producing this information and an appropriate locus for the activity would need to be found.

As things stand, it appears that legislative agencies will be asked to devote more and more of their resources to policy formulation work. If this should continue, the balance of analytical power -- as between the executive and legislative branches of government -- could eventually be transformed, unless there is a concomitant development of a similar resource in the executive branch. Since this would leave the executive branch at a significant disadvantage as it proposes new programs to implement Administration policy, it seems reasonable to expect that, over the long term, more careful evaluative scrutiny by the legislative branch should help produce improved policy formulation in government generally. This is an outcome devoutly to be wished.

References

- Auditor General of Canada, Program Evaluation Report of the Auditor General, Chapter 3, 1983, Government of Canada, Ottawa.
- ERS Standards Committee, "Evaluation Research Society Standards for Program Evaluation," New Directions for Program Evaluation, Standards for Evaluation Practice, Peter H. Rossi, Editor, September 1982 (Jossey-Bass, Inc.).
- Hoaglin, David C., Richard J. Light, Bucknam McPeck, Frederick Mosteller, and Michael A. Stoto, Data for Decisions, Abt Books, Cambridge, Mass., 1982.
- Rodney Nichols, "Mission-Oriented R&D," Science, Vol. 172, pages 29-36, 2 April 1971.